

# A Degenerate Isoperimetric Problem in the Plane

Jiri Dadok<sup>\*1</sup> and Peter Sternberg<sup>†2</sup>

<sup>1,2</sup>*Department of Mathematics, Indiana University, Bloomington, IN 47405*

July 29, 2016

**Abstract:** We establish sufficient conditions for existence of curves minimizing length as measured with respect to a degenerate metric on the plane while enclosing a specified amount of Euclidean area. Non-existence of minimizers can occur and examples are provided. This continues the investigation begun in [1] where the metric  $ds^2$  near the singularities equals a quadratic polynomial times the standard metric. Here we allow the conformal factor to be any smooth non-negative potential vanishing at isolated points *provided* the Hessian at these points is positive definite. These isoperimetric curves, appropriately parametrized, arise as traveling wave solutions to a bi-stable Hamiltonian system.

## 1 Introduction

Given a complete metric on  $\mathbb{R}^2$  conformal to the standard Euclidean metric, it is not hard to show that for any two points  $p$  and  $q$  in  $\mathbb{R}^2$  and any positive number  $A$ , one can find a shortest curve  $\gamma$  joining  $p$  and  $q$  among all competitors which, together with a distance minimizing geodesic  $\gamma_0$  from  $q$  to  $p$  enclose a specified amount of *Euclidean* area. This variant of an isoperimetric problem, in which length and area are measured with respect to different metrics, becomes more subtle, however, when the conformal factor,  $F$ , is allowed to vanish at one or more points in the plane.

Motivated by the observation that such isoperimetric curves can be used to build traveling wave solutions to certain Hamiltonian systems, the authors of [1] analyze this problem for the case where  $W := F^2$  is a non-negative map from  $\mathbb{R}^2$  to  $\mathbb{R}$  vanishing at two points, say  $\mathbf{p}_-$  and  $\mathbf{p}_+$ , and where the isoperimetric curve is required to join these two potential wells. Thus, the problem takes the form

$$\inf \left\{ \int_a^b F(\gamma(t)) |\gamma'(t)| dt : \gamma(a) = \mathbf{p}_-, \gamma(b) = \mathbf{p}_+, \int_\gamma \omega_0 = A \right\}, \quad (1.1)$$

where  $\omega_0$  is the 1-form  $p_1 dp_2$ , and so the interpretation as a Euclidean area constraint comes through an application of Stokes Theorem to the closed curve  $\gamma \cup \gamma_0$  in that  $d\omega_0 = dp_1 dp_2$ . An

---

<sup>\*</sup>dadok@indiana.edu

<sup>†</sup>sternber@indiana.edu

equivalent version of the problem results from replacing  $\omega_0$  by any 1-form  $\omega$  such that  $d\omega = dp_1 dp_2$ .

What makes this particular isoperimetric problem non-standard is both the degeneracy of the conformal factor and the fact that length is measured with respect to a metric given by  $F$  while area is measured with respect to the Euclidean metric. There is a vast literature on isoperimetric problems and in particular isoperimetric curves with assorted assumptions on the conformal factor or “density,” though to our knowledge none address this combination of degeneracy and mixture of metrics. We mention, for example, [3, 4, 5, 7, 9] but of course there are many others.

Under the assumption that  $W$  is exactly equal to a nondegenerate quadratic polynomial in neighborhoods of the zeros  $\mathbf{p}_\pm$ , such an isoperimetric curve joining  $\mathbf{p}_-$  to  $\mathbf{p}_+$  is shown to exist in [1] for any amount of Euclidean area  $A$  via a calibration argument. Near the wells, these minimizers are realized as integral curves of an explicit vector field and interestingly one finds that they often spiral into the zeros of  $F$ . On the other hand, if  $F(p)/|p - \mathbf{p}_\pm| \rightarrow 0$  as  $|p - \mathbf{p}_\pm| \rightarrow 0$  for either zero, then minimizers do not exist for any  $A \neq \int_{\gamma_0} \omega_0$ .

In this article we relax the rather stringent assumption that  $W$  is a quadratic near its isolated zeros and instead simply assume that it is smooth and has positive definite Hessian matrix  $D^2W$  at its two zeros, along with an assumption on behavior at infinity to ensure completeness of the metric. Since near either potential well  $W$  is well-approximated by its quadratic Taylor polynomial it would seem reasonable to conjecture that the same type of existence result as the one found in [1] is true in this more general setting. Surprisingly this turns out to be false.

Our main result, Theorem 5.3, states that when  $A$  is sufficiently close to  $\int_{\gamma_0} \omega_0$ , an isoperimetric curve exists. However, in Section 4 we reveal through some counter-examples that if  $|A - \int_{\gamma_0} \omega_0|$  is not small, a minimizer may fail to exist. The mechanism for non-existence is associated with minimizing sequences packing part of the area around the wells. In the limit this part ‘disappears’ and the limiting curve encloses a smaller amount of area.

The existence result relies on the introduction of a complete Carnot-Carathéodory length metric on  $\mathbb{R}^3$  (cf. [6]) with singularities over the wells. The geodesics for this metric project to the plane as solutions to our problem. This re-phrasing of the problem is explained in Section 2 but roughly speaking, the “extra” component of competing curves keeps track of the amount of Euclidean area accumulated by a planar curve as it journeys from a point  $p$  to a point  $q$ .

We review in Section 3 the results from [1] on the case of quadratic  $W$ , along with some essential extensions. Section 4 contains the counter-examples, Section 5 contains the proof of the main existence result and in Section 6 we recall from [1] how under further non-degeneracy assumptions, these isoperimetric curves can serve as traveling wave solutions to the Hamiltonian system

$$\mathbb{J}u_t = \Delta u - \nabla W_u(u) \quad \text{for } u : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^2 \quad (1.2)$$

where  $\mathbb{J}$  denotes the symplectic matrix given by

$$\mathbb{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This not yet well-understood system of partial differential equations represents a form of conservative dynamics associated with the well-studied vector Allen-Cahn (or vector Modica-Mortola) energy

$$u \mapsto \int \frac{1}{2} |\nabla u|^2 + W(u),$$

and could provide an interesting model for phase transitions in the Hamiltonian setting. One suspects that having a solid understanding of traveling wave solutions will be a helpful first step in subsequent analysis of the dynamics. Invoking the traveling wave ansatz

$$u(x_1, \dots, x_n, t) = U(x_1 - vt)$$

in (1.2), one sees that  $U : \mathbb{R} \rightarrow \mathbb{R}^2$  must solve the system of ODE's

$$-v\mathbb{J}U' = U'' - \nabla_u W(U) \quad \text{for } -\infty < y < \infty, \quad U(\pm\infty) = \mathbf{p}_\pm. \quad (1.3)$$

It turns out that isoperimetric curves  $\gamma$  solving (1.1), appropriately parametrized, will solve this system where the wave speed  $v$  depends on the area constraint  $A$  and can also be related to the geodesic curvature  $\kappa_g$  of  $\gamma$  through the formula  $v = \kappa_g W(\gamma)$ , a quantity that turns out to be constant along the isoperimetric curves, cf. Remark 2.7.

**Acknowledgment.** The authors wish to thank Stan Alama, Lia Bronsard and Andres Contreras for many helpful conversations on this problem. P.S. also wishes to acknowledge the support of the National Science Foundation through D.M.S. 1362879.

## 2 A re-casting of the problem in terms of curves in 3-space

We begin by restating our hypotheses more precisely. The arguments we will present apply to the setting where the conformal factor vanishes at any set of isolated points but for ease of presentation we will take the zero set to be two points. Thus, we assume throughout that  $W : \mathbb{R}^2 \rightarrow [0, \infty)$  is a smooth function vanishing only at two points  $\mathbf{p}_-$  and  $\mathbf{p}_+$  with positive definite Hessian matrix  $D^2W(\mathbf{p}_\pm) > 0$ . We also assume

$$\liminf_{|p| \rightarrow \infty} W(p) > 0.$$

Then we define  $F : \mathbb{R}^2 \rightarrow [0, \infty)$  by

$$F(p) := \sqrt{W(p)}$$

We first recall from [10] the following result (see also [8]):

**Theorem 2.1.** *For any two points  $p$  and  $q$  in  $\mathbb{R}^2$  there exists a (not necessarily unique) minimizer to the problem*

$$\bar{d}_F(p, q) := \inf \left\{ E(\gamma) : \gamma : [a, b] \rightarrow \mathbb{R}^2 \text{ is locally Lipschitz continuous and } \gamma(a) = p, \gamma(b) = q \right\}. \quad (2.1)$$

For the case  $p = \mathbf{p}_-$  and  $q = \mathbf{p}_+$  we choose  $\gamma_0$  as a minimizer to (2.1) and we refer to such a curve as an *unconstrained planar geodesic* to distinguish it from curves solving the constrained problem (2.3) below and from certain geodesics in three dimensions to be introduced shortly.

For any locally Lipschitz continuous mapping  $\gamma : [a, b] \rightarrow \mathbb{R}^2$ , writing  $\gamma(t) = (\gamma^{(1)}(t), \gamma^{(2)}(t))$  we now define

$$E(\gamma) := \int_a^b F(\gamma(t)) |\gamma'(t)| dt \quad \text{and} \quad \mathcal{A}(\gamma) := \int_\gamma \omega_0 = \int_a^b \gamma^{(1)}(t) \gamma^{(2)'}(t) dt, \quad (2.2)$$

where  $\omega_0 = p_1 dp_2$ . Our goal is to establish the existence of a solution to the problem

$$m_A := \inf \left\{ E(\gamma) : \gamma : [a, b] \rightarrow \mathbb{R}^2 \text{ is locally Lipschitz continuous, } \gamma(a) = \mathbf{p}_-, \gamma(b) = \mathbf{p}_+, \mathcal{A}(\gamma) = A \right\} \quad (2.3)$$

for certain values of the area constraint  $A$ .

It will be convenient in what follows to introduce notation for the projection from  $\mathbb{R}^3$  down to  $\mathbb{R}^2$  so we will write  $\Pi(P) := (p_1, p_2)$  for any point  $P = (p_1, p_2, p_3) \in \mathbb{R}^3$ .

We also introduce a certain lifting of a planar curve into  $\mathbb{R}^3$ . To any locally Lipschitz continuous planar curve  $\gamma : [a, b] \rightarrow \mathbb{R}^2$  we will frequently associate a locally Lipschitz curve in 3-space  $\Gamma : [a, b] \rightarrow \mathbb{R}^3$  via the properties  $\Pi(\Gamma) = \gamma$  so that  $\Gamma^{(j)} = \gamma^{(j)}$  for  $j = 1, 2$  and

$$\Gamma^{(3)'}(t) = \gamma^{(1)}(t) \gamma^{(2)'}(t) \text{ for a.e. } t \in (a, b). \quad (2.4)$$

We point out that (2.4) only determines the third component function of the curve  $\Gamma$  up to an arbitrary additive constant and that

$$\Gamma^{(3)}(b) - \Gamma^{(3)}(a) = \mathcal{A}(\gamma).$$

With this notation in hand we define the singular Carnot-Carathéodory length metric  $d_F : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, \infty)$  (cf. [6]) via

$$d_F(P, Q) := \inf \left\{ E(\gamma) : \Pi(\Gamma) = \gamma, \Gamma(a) = P, \Gamma(b) = Q, \Gamma \text{ satisfies (2.4)} \right\}. \quad (2.5)$$

From now on, when a curve in three-dimensions satisfies (2.4) we will say it is *admissible* in (2.5) and we comment that for any two points  $P = (p_1, p_2, p_3)$  and  $Q = (q_1, q_2, q_3)$ , the quantity  $d_F(P, Q)$  only depends on  $p_3$  and  $q_3$  through the difference  $q_3 - p_3$ . One can also readily check through the association of two-dimensional curves and admissible three-dimensional curves via lifting that an equivalent characterization of  $d_F$  is

$$d_F(P, Q) = \inf \left\{ E(\gamma) : \gamma(a) = (p_1, p_2), \gamma(b) = (q_1, q_2), \int_a^b \gamma^{(1)}(t) \gamma^{(2)'}(t) dt = q_3 - p_3 \right\}. \quad (2.6)$$

One sees immediately from a comparison with (2.3) that when  $\Pi(P) = \mathbf{p}_-$ ,  $\Pi(Q) = \mathbf{p}_+$  and  $q_3 - p_3 = A$ , one has

$$d_F(P, Q) = m_A.$$

The advantage of this re-formulation of problem (2.3) comes through the introduction of the notion of length for *any* curve  $\Gamma : [a, b] \rightarrow \mathbb{R}^3$ , not necessarily one satisfying (2.4), via

$$L(\Gamma) := \sup_{\{t_j\}_{j=1}^N \in \mathcal{P}([a, b])} \sum_{j=1}^N d_F(\Gamma(t_j), \Gamma(t_{j+1})), \quad (2.7)$$

where  $\mathcal{P}([a, b])$  is the set of finite partitions of  $[a, b]$ . As we will see below, there *always* exists a curve  $\Gamma$  in  $\mathbb{R}^3$  of least length—though not necessarily admissible!— with endpoints  $P$  and  $Q$  satisfying the property  $L(\Gamma) = d(P, Q)$  whereas solutions to (2.3) sometimes do not exist.

To understand better the relation between the metric  $d_F$  and the length of curves we first observe:

**Proposition 2.2.** *The metric  $d_F$  given in (2.5) defines a length metric on  $\mathbb{R}^3$  (cf. [6]); that is, for any points  $P$  and  $Q$  in  $\mathbb{R}^3$  one has*

$$d_F(P, Q) = \inf_{\Gamma} L(\Gamma) \quad (2.8)$$

where the infimum on the right is taken over all curves  $\Gamma : [a, b] \rightarrow \mathbb{R}^3$  that are Lipschitz continuous with respect to the metric  $d$  and satisfy  $\Gamma(a) = P$ ,  $\Gamma(b) = Q$  and  $\Gamma^{(3)'}(t) = \Gamma^{(1)'}(t)\Gamma^{(2)'}(t)$  for a.e.  $t \in (a, b)$ , thus making the metric space into a length space, cf. [2].

*Proof.* Away from any zeros of  $F$  the metric property of  $d_F$  is standard but there is one subtlety to check: Given two distinct points  $P$  and  $Q$  that both project to either  $\mathbf{p}_-$  or  $\mathbf{p}_+$ , we must verify that  $d_F(P, Q) \neq 0$ . This follows from the assumption of non-degeneracy of the Hessian matrices  $D^2W(\mathbf{p}_{\pm}) > 0$ , since locally, for instance near  $\mathbf{p}_-$ , one can fit a non-degenerate quadratic polynomial under  $W$ , say  $\tilde{W}(p) = \tilde{\lambda}_1^2 p_1^2 + \tilde{\lambda}_2^2 p_2^2$ , so that

$$F(p) := \sqrt{W(p)} \geq \tilde{F}(p) := \sqrt{\tilde{\lambda}_1^2 p_1^2 + \tilde{\lambda}_2^2 p_2^2}.$$

It follows that  $d_F(P, Q) \geq d_{\tilde{F}}(P, Q)$  and the behavior of the metric for such homogeneous conformal factors was completely worked out in [1]. The relevant results are all reviewed in Section 3 to follow but in particular through an application of Corollary 3.4 we have that

$$d_F(P, Q) = d_F((\mathbf{p}_-, 0), (\mathbf{p}_-, q_3 - p_3)) \geq d_{\tilde{F}}((\mathbf{p}_-, 0), (\mathbf{p}_-, q_3 - p_3)) = (\tilde{\lambda}_1 + \tilde{\lambda}_2) |q_3 - p_3| > 0$$

in case  $P$  and  $Q$  both project down to  $\mathbf{p}_-$  with a similar inequality holding if  $\Pi(P) = \Pi(Q) = \mathbf{p}_+$ .

To establish the equivalence (2.8) we can use the fact that

$$L(\Gamma) = E(\Pi(\Gamma)) \quad (2.9)$$

for any admissible  $\Gamma$  which follows by the same argument used to prove Theorem 2.5 of [10]. Taking the infimum of both sides leads to the result.  $\square$

**Remark 2.3.** *We wish to emphasize the crucial role that the non-degeneracy of the Hessian matrices  $D^2W(\mathbf{p}_{\pm})$  plays here. For example, suppose  $\mathbf{p}_- = (0, 0)$  and in a neighborhood of the origin  $W$  takes the form of a radial function  $W = W(r) = r^\alpha$  for  $\alpha > 2$ . Then take  $P = (0, 0, 0)$  and  $Q = (0, 0, A)$  for any  $A > 0$ . With  $r_j$  chosen so that  $j\pi r_j^2 = A$ , the sequence of planar curves  $\{\gamma_j\}$  consisting of the directed line segment  $\ell_j$  joining  $(0, 0)$  to  $(0, r_j)$ , followed by  $j$  circles of radius  $r_j$  and ending with  $-\ell_j$  returning to the origin has the property that  $E(\gamma_j) \rightarrow 0$  while  $\mathcal{A}(\gamma_j) = A$ . In light of (2.6), this shows that  $d_F((0, 0, 0), (0, 0, A)) = 0$  when such a degeneracy is allowed and so we lose the metric property. See Section 2.2 of [1] for more discussion.*

The advantage of working in the length space context is the following Hopf-Rinow type result on existence of geodesics. The proof involves a simple application of the Arzela-Ascoli Theorem working in the metric space of  $\mathbb{R}^3$  endowed with the  $d_F$  metric.

**Theorem 2.4.** ([2], Prop. 3.7) *For every pair of points  $P$  and  $Q$  in  $\mathbb{R}^3$  there exists a minimizing geodesic  $\Gamma : [a, b] \rightarrow \mathbb{R}^3$  joining  $P$  to  $Q$  in the sense that  $\Gamma(a) = P$ ,  $\Gamma(b) = Q$  and  $L(\Gamma) = d(P, Q)$ . This geodesic, which may not be admissible in the sense of (2.4), can be realized as the uniform limit in the metric  $d_F$  of a minimizing sequence of admissible curves as described in Proposition 2.2.*

Whenever this geodesic contains no vertical line segments over the wells, that is, whenever it projects down to a curve, or union of curves, each of which avoids the wells  $\mathbf{p}_-$  and  $\mathbf{p}_+$  except perhaps at their endpoints, then it is elementary to check that the projection minimizes  $E$  subject to fixed endpoints and the appropriate constraint value  $q_3 - p_3$ , and this is the content of the next proposition. On the other hand, we emphasize that it is *not* the case that such a geodesic projects to a constrained minimizer of  $E$  if it *does* include a vertical line segment over either of the wells, since in this case the projection has ‘lost area’ and fails to satisfy the constraint value.

**Proposition 2.5.** *For any two points  $P$  and  $Q$  in  $\mathbb{R}^3$ , let  $\Gamma : [0, 1] \rightarrow \mathbb{R}^3$  denote a geodesic joining them, as guaranteed by Theorem 2.4. Suppose the closed set  $\{t \in [0, 1] : \Pi(\Gamma(t)) \in \{\mathbf{p}_-, \mathbf{p}_+\}\}$  contains no intervals. Then  $\gamma := \Pi(\Gamma)$  minimizes  $E$  among all locally Lipschitz curves  $\zeta$  joining  $\Pi(P)$  to  $\Pi(Q)$  and satisfying  $\mathcal{A}(\zeta) = q_3 - p_3$ . In particular, if the vertical fibers over  $\mathbf{p}_-$  and  $\mathbf{p}_+$  are not geodesics then a solution to (2.3) exists for all values of  $A$ .*

*Proof.* Consider the restriction of  $\Gamma$  to a parameter interval  $(a, b) \subset [0, 1]$  such that the projection of  $\Gamma(a, b)$  misses the wells  $\mathbf{p}_-$  and  $\mathbf{p}_+$ . Recall from Theorem 2.4 that  $\Gamma$  can be realized as the uniform limit with respect to the metric  $d$  of a minimizing sequence of admissible curves  $\Gamma_j : [0, 1] \rightarrow \mathbb{R}^3$  for  $j = 1, 2, \dots$  parametrized with constant speed, i.e. curves satisfying (2.4) as well as the conditions

$$E(\gamma_j) \rightarrow d_F(P, Q) \text{ as } j \rightarrow \infty \quad \text{and} \quad F(\gamma_j(t)) |\gamma_j'(t)| = E(\gamma_j) \text{ for every } j, \quad (2.10)$$

where  $\gamma_j := \Pi(\Gamma_j)$ .

The assumption that the projection  $\gamma(a, b)$  misses the points  $\mathbf{p}_+$  and  $\mathbf{p}_-$  implies that for any  $\delta > 0$ , there exists a positive constant  $C_\delta$  such that for all  $j$  large and for all  $t \in [a + \delta, b - \delta]$  one has  $F(\gamma_j(t)) \geq C_\delta$ . Then applying (2.10) we find that  $\sup_{t \in [a + \delta, b - \delta]} |\gamma_j'(t)|$  is bounded by a constant independent of  $j$ . Consequently we can apply Arzela-Ascoli to pass to a subsequence (still denoted by  $\gamma_j$ ) such that  $\gamma_j \rightarrow \gamma$  on  $[a + \delta, b - \delta]$  uniformly in the Euclidean metric, and such that

$$\gamma_j' \rightharpoonup \gamma' \text{ weakly in } L^2((a + \delta, b - \delta)).$$

Then for any  $t_1, t_2 \in [a + \delta, b - \delta]$  it follows that

$$\int_{t_1}^{t_2} \gamma^{(1)} \gamma^{(2)'} dt = \lim_{j \rightarrow \infty} \int_{t_1}^{t_2} \gamma_j^{(1)} \gamma_j^{(2)'} dt = \lim_{j \rightarrow \infty} \int_{t_1}^{t_2} \Gamma_j^{(3)'} dt = \Gamma^{(3)}(t_2) - \Gamma^{(3)}(t_1).$$

Since  $\delta$  is arbitrary, we conclude that  $\Gamma^{(3)'}(t) = \gamma^{(1)}(t) \gamma^{(2)'}(t)$  for a.e.  $t \in (a, b)$ , that is,  $\Gamma$  is admissible in the sense of (2.4). In particular, we note that admissibility precludes the possibility that  $\Gamma$  contains any vertical line segments over points in  $\mathbb{R}^2$  other than  $\mathbf{p}_-$  or  $\mathbf{p}_+$ . We then invoke the property of length-minimizing geodesics that for any sub-interval  $[t_1, t_2] \subset [0, 1]$ , one has that  $L(\Gamma|_{[t_1, t_2]})$  minimizes length when compared to any other curve joining  $\Gamma(t_1)$  to  $\Gamma(t_2)$ . In light

of (2.9), this means that the projection  $\gamma|_{[a,b]}$  minimizes  $E$  among competitors  $\zeta : [a, b] \rightarrow \mathbb{R}^2$  joining  $\gamma(a)$  to  $\gamma(b)$  and sharing the same constraint value  $\mathcal{A}(\zeta) = \mathcal{A}(\gamma|_{[a,b]})$ . Thus,

$$d_F(\Gamma(a), \Gamma(b)) = L(\Gamma|_{[a,b]}) = E(\gamma|_{[a,b]}). \quad (2.11)$$

Writing  $\Gamma$  as a union of such ‘good’ sub-arcs and summing over this union we conclude that (2.11) holds with  $[a, b]$  replaced by  $[0, 1]$ .

Finally, if the metric  $d_F$  is such that the vertical fibers over  $\mathbf{p}_-$  and  $\mathbf{p}_+$  are not geodesic then  $\Gamma$  always projects to a solution of (2.3) for any  $A$ -values since this condition would preclude the possibility that  $\Gamma$  contains any vertical segments over the wells. An example of this phenomenon is given in Section 4.  $\square$

**Corollary 2.6.** *For any two points  $P$  and  $Q$  in  $\mathbb{R}^3$ , let  $\Gamma : [0, 1] \rightarrow \mathbb{R}^3$  denote a minimizing geodesic joining them, as guaranteed by Theorem 2.4. Let  $[a, b] \subset [0, 1]$  be any interval such that the projection  $\gamma := \Pi(\Gamma)$  avoids  $\mathbf{p}_+$  and  $\mathbf{p}_-$  for all  $t \in (a, b)$ . Then  $\Gamma$  restricted to any such interval  $(a, b)$  is smooth, satisfies the relation (2.4), and has a smooth projection that satisfies the system of ODE’s*

$$-\left(F(\gamma) \frac{\gamma'}{|\gamma'|}\right)' + |\gamma'| \nabla F(\gamma) = -\lambda (\gamma')^\perp \quad (2.12)$$

for some constant  $\lambda$ . Here  $(x, y)^\perp := (-y, x)$ . If  $\Gamma$  is expressible as a union of such curves, that is, if  $\Gamma$  contains no vertical line segments over  $\mathbf{p}_-$  or  $\mathbf{p}_+$ , then (2.12) holds along each curve in the union with the same value of  $\lambda$ .

*Proof.* The argument that  $\Gamma$  is admissible along the parameter interval  $(a, b)$  follows as in the proof of Theorem 2.5. Hence, as in that proof the length-minimizing property of  $\Gamma$  restricted to  $[a, b]$  implies the area-constrained  $E$ -minimality of the projection  $\gamma|_{[a,b]}$ . In other words, we know that

$$\int_a^b F(\gamma) |\gamma'| dt \leq \int_a^b F(\zeta) |\zeta'| dt$$

whenever  $\zeta : [a, b] \rightarrow \mathbb{R}^2$  satisfies  $\zeta(a) = \gamma(a)$ ,  $\zeta(b) = \gamma(b)$  and

$$\int_a^b \zeta^{(1)} \zeta^{(2)'} dt = \int_a^b \gamma^{(1)} \gamma^{(2)'} dt.$$

Applying the theory of Lagrange multipliers we conclude that  $\gamma$  restricted to  $[a, b]$  must satisfy the criticality condition

$$\delta E(\gamma; \tilde{\gamma}) = \lambda \delta \mathcal{A}(\gamma; \tilde{\gamma}) \quad \text{for any compactly supported variation } \tilde{\gamma}$$

where  $\lambda \in \mathbb{R}$ . (Here  $\delta$  refers to first variation.) Since  $\delta \mathcal{A}(\gamma; \tilde{\gamma}) = -\int (\gamma')^\perp \cdot \tilde{\gamma}$  and since we know that for  $t \in (a + \delta, b - \delta)$  one has  $F(\gamma) \geq C_\delta > 0$ , we can apply standard regularity theory for ODE’s to conclude that  $\gamma$  is smooth and classically solves (2.12) through a routine computation of  $\delta E(\gamma; \tilde{\gamma})$ . Necessarily the lifting of  $\gamma$  to  $\Gamma$  via (2.4) over this sub-interval is then also smooth.

Finally we note that if  $\Gamma$  consists of more than one such arc then the Lagrange multiplier  $\lambda$  cannot vary from arc to arc for if it did, one could create a variation through a bump on one arc and a compensating bump on the other arc that preserved the total constraint value and violated the condition of vanishing first variation.  $\square$

**Remark 2.7. Characterization of criticality in terms of geodesic curvature.** A more geometric way to view the condition of criticality for a planar curve  $\gamma$  is the following. Suppose we parametrize  $\gamma$  by degenerate arclength, say  $\ell$ , so that

$$\|\gamma'(\ell)\|_g = \langle \gamma'(\ell), \gamma'(\ell) \rangle_g^{1/2} := F(\gamma(\ell)) |\gamma'(\ell)| = 1 \text{ for some parameter interval } 0 < \ell < L,$$

where our metric  $g$  is given by  $g_{ij} = F^2 \delta_{ij}$  in standard coordinates. Now we consider the flow given by the exponential map  $\tilde{\gamma}(\ell, \tau) = \exp_{\gamma(\ell)} \tau V$  where  $V$  is a normal vector field given by say  $f N$  with  $f : [0, L] \rightarrow \mathbb{R}$  and  $\|N\|_g = 1$  so that

$$N(\ell) = \frac{1}{F(\gamma(\ell))} \frac{(\gamma'(\ell))^\perp}{|\gamma'(\ell)|} = \frac{1}{F(\gamma(\ell))} n(\ell),$$

with  $n$  denoting the unit normal in the Euclidean metric. Then we can carry out the first variation calculation as

$$\begin{aligned} \frac{d}{d\tau} E(\tilde{\gamma}) \Big|_{\tau=0} &= \frac{d}{d\tau} \int_0^L \left\langle \frac{\partial \tilde{\gamma}}{\partial \ell}, \frac{\partial \tilde{\gamma}}{\partial \ell} \right\rangle_g^{1/2} d\ell \\ &= \int_0^L \left\langle \nabla_{\gamma'} \gamma', V \right\rangle_g^{1/2} d\ell = \int_0^L \kappa_g f d\ell, \end{aligned} \quad (2.13)$$

where  $\kappa_g$  denotes the geodesic curvature of  $\gamma$  with respect to the metric  $g$ . To restrict the flow to variations that preserve Euclidean area to leading order we note that  $V = \frac{f}{F} n$  so the allowable  $f$  are those such that

$$0 = \int_0^{s(L)} \frac{f}{F} ds = \int_0^L \frac{f}{F^2} d\ell,$$

where  $s$  denotes Euclidean arclength. Finally, letting  $h(\ell) := \frac{f(\ell)}{F^2(\gamma(\ell))}$  we see from (2.13) that criticality with a Euclidean area constraint means that

$$\int_0^L F^2(\gamma(\ell)) \kappa_g(\ell) h(\ell) d\ell = 0 \text{ for all } h : [0, L] \rightarrow \mathbb{R} \text{ such that } \int_0^L h(\ell) d\ell = 0.$$

We conclude that  $F^2 \kappa_g = \lambda$ , a constant, along  $\gamma$ . Of course in minimizing length with fixed area, one expects curvature to arise in the criticality condition but the point here is that the mis-match of metrics leads to the extra factor of  $F^2$ . Relating this to (2.12) we can identify an expression for geodesic curvature as

$$\kappa_g = \frac{1}{F^2(\gamma)} \left[ \frac{1}{|\gamma'|^2} \left( F(\gamma) \frac{\gamma'}{|\gamma'|} \right)' - \frac{1}{|\gamma'|} \nabla F(\gamma) \right] \cdot (\gamma')^\perp$$

where  $\gamma$  is given by an arbitrary parametrization.

### 3 The case of quadratic $W$

In [1] the problem (2.3) is solved for the case where  $W$  is given by a non-degenerate quadratic in a neighborhood of  $\mathbf{p}_-$  and  $\mathbf{p}_+$ . The key there is to first find the solution for the case where



$W$  vanishes only at one well, taken for convenience to be the origin. Choosing coordinate axes given by the eigenvectors of the Hessian matrix  $D^2W(0, 0)$ , one considers  $W (= F^2)$  to be of the form

$$W(p_1, p_2) = \lambda_1^2 p_1^2 + \lambda_2^2 p_2^2 \quad (3.1)$$

for positive constants  $\lambda_1$  and  $\lambda_2$ . For this special case where  $W$  is purely quadratic so that  $\sqrt{W}$  is homogeneous of degree one, we will introduce the notation  $F_H$  via

$$F_H(p) := \sqrt{\lambda_1^2 p_1^2 + \lambda_2^2 p_2^2}. \quad (3.2)$$

For the rest of the article, when we compute distance using the metric  $d_{F_H}$  defined by replacing a general conformal factor  $F$  by  $F_H$  in (2.5) we will refer to this metric as the *homogeneous metric*.

Then the following result is proven:

**Theorem 3.1.** (cf. Thm 2.5 of [1]). *For any value  $A \in \mathbb{R}$  and any point  $p_0 \in \mathbb{R}^2 \setminus \{(0, 0)\}$  there exists a unique solution to*

$$\inf \left\{ E_H(\gamma) : \gamma : [a, b] \rightarrow \mathbb{R}^2 \text{ is locally Lipschitz continuous, } \gamma(a) = p_0, \gamma(b) = (0, 0) \text{ and } \mathcal{A}(\gamma) = A \right\}, \quad (3.3)$$

where

$$E_H(\gamma) := \int_a^b F_H(\gamma) |\gamma'| dt.$$

This minimizer, denoted by  $\gamma_\beta$ , is given explicitly as the integral curve of the vector field

$$V_\beta(p) := (\cos \beta) \Theta(p) - (\sin \beta) R(p)$$

that joins  $p_0$  to the origin where

$$R(p) := \frac{\nabla \tilde{r}(p)}{F_H(p)^2}, \quad \text{with} \quad \tilde{r}(p) := \frac{1}{2}(\lambda_1 p_1^2 + \lambda_2 p_2^2), \quad (3.4)$$

$$\text{and} \quad \Theta(p) := \frac{(-\lambda_2 p_2, \lambda_1 p_1)}{F_H(p)^2}.$$

Here  $\beta$  is selected so that

$$\frac{\tilde{r}(p_0)}{\lambda_1 + \lambda_2} \cot \beta = A + C_0, \quad (3.5)$$

where  $C_0$  is an explicit constant depending only on  $p_0$  such that  $C(p_0) \rightarrow 0$  as  $|p_0| \rightarrow 0$ . In particular, the planar curve  $\gamma_\beta$  avoids the origin except at the endpoint  $t = b$  and satisfies a Euclidean arclength bound of the form

$$\int_a^b |\gamma'_\beta| dt < C(\lambda_1, \lambda_2, p_0, A). \quad (3.6)$$

**Corollary 3.2.** *For  $F = F_H = \sqrt{W}$  with  $W$  given by (3.1), and for  $P = (p_0, 0)$  and  $Q = (0, 0, A)$  with  $p_0 \neq (0, 0)$  the unique minimizing geodesic  $\Gamma_\beta = (\Gamma_\beta^{(1)}, \Gamma_\beta^{(2)}, \Gamma_\beta^{(3)})$  joining  $P$  to  $Q$  is given by*

$$(\Gamma_\beta^{(1)}, \Gamma_\beta^{(2)}) = \gamma_\beta \quad \text{and} \quad \Gamma_\beta^{(3)'} = \gamma_\beta^{(1)} \gamma_\beta^{(2)'}, \quad \Gamma_\beta^{(3)}(a) = 0.$$

Furthermore, one has the identities

$$L(\Gamma_\beta) = E(\gamma_\beta) = d_{F_H}(P, Q) = \tilde{r}(p_0) \csc \beta \quad (3.7)$$

with the angle  $\beta \in (-\pi/2, \pi/2) \setminus \{0\}$  depending on  $\mathbf{p}_0$  and  $A$  through (3.5).

Lastly, if we parametrize  $\Gamma$  by  $A$  rather than  $\beta$ , one has

$$\frac{dL(\Gamma_A)}{dA} = (\lambda_1 + \lambda_2) \cos \beta_A \quad \text{where } \beta_A \text{ is determined by (3.5).} \quad (3.8)$$

*Proof.* This follows immediately from the fact that (2.9) holds for admissible curves which implies that whenever there exists an admissible planar curve that minimizes  $E$  then its lifting minimizes the three-dimensional length  $L$ . The identity (3.7), phrased in terms of  $E(\gamma_\beta)$  is derived within the proof of Thm 2.5 of [1]. Then (3.8) follows by direct calculation using (3.5) and (3.7).  $\square$

The proof of Theorem 3.1 follows by a calibration argument and though it was not needed and so did not appear in [1], we note here that even in the limiting case where  $\beta = 0$  so that  $V_\beta(p) = V_0(p) = \Theta(p)$ , the calibration approach still works. The result is an identification of certain ellipses centered at the origin as closed isoperimetric curves for the homogeneous metric addressed in this section.

**Corollary 3.3.** *For any  $p_0 \in \mathbb{R}^2 \setminus \{0\}$  the ellipse  $\tilde{r}(p) = \tilde{r}(p_0)$  minimizes  $E_H$  among all competing closed curves passing through  $p_0$  and enclosing the same amount of Euclidean area  $A$ . Furthermore, the value of  $E_H$  for this minimizing ellipse is given by  $(\lambda_1 + \lambda_2)A$ .*

*Proof.* The calibration argument for the Corollary is as follows. Define the 1-form

$$\omega := -\frac{\lambda_2}{\lambda_1 + \lambda_2} p_2 dp_1 + \frac{\lambda_1}{\lambda_1 + \lambda_2} p_1 dp_2,$$

and note that  $d\omega$  is just the Euclidean area form  $dp_1 dp_2$ . Parametrize the ellipse by degenerate arclength, say  $\bar{\gamma} : [0, L_{\bar{\gamma}}] \rightarrow \mathbb{R}^2$ , by expressing it as the integral curve of the vector field  $V_0$  so that

$$\bar{\gamma}' = V_0(\bar{\gamma}), \quad \bar{\gamma}(0) = p_0 \quad \text{where} \quad V_0(p_1, p_2) := \frac{(-\lambda_2 p_2, \lambda_1 p_1)}{F_H(p)^2}$$

and we have that  $F(\bar{\gamma}(\ell)) |\bar{\gamma}'|(\ell) = 1$ . Then let  $\gamma : [0, L_\gamma] \rightarrow \mathbb{R}^2$  be any other closed curve parametrized by degenerate arclength with  $\gamma(0) = p_0$  that satisfies the same area constraint. Then by Stokes Theorem we have  $\mathcal{A}(\gamma) := \int_\gamma p_1 dp_2 = \int_\gamma \omega = \int_{\bar{\gamma}} \omega$ . Finally we compute that

$$(\lambda_1 + \lambda_2) \int_{\bar{\gamma}} \omega = \int_0^{L_{\bar{\gamma}}} (\lambda_2 \bar{\gamma}^{(2)}, \lambda_1 \bar{\gamma}^{(1)}) \cdot \bar{\gamma}' d\ell = \int_0^{L_{\bar{\gamma}}} 1 d\ell = L_{\bar{\gamma}} = E(\bar{\gamma})$$

while

$$(\lambda_1 + \lambda_2) \int_\gamma \omega < \int_0^{L_\gamma} |(\lambda_2 \gamma^{(2)}, \lambda_1 \gamma^{(1)})| \cdot |\gamma'| d\ell = \int_0^{L_\gamma} 1 d\ell = L_\gamma = E(\gamma).$$

Since  $\int_{\bar{\gamma}} \omega$  is the amount of Euclidean area  $A$  enclosed by the ellipse  $\bar{\gamma}$  we are done.  $\square$

**Corollary 3.4.** *For any  $A \in \mathbb{R}$  one has the identity*

$$d_{F_H}((0, 0, 0), (0, 0, A)) = (\lambda_1 + \lambda_2) |A|.$$

*Proof.* For any  $\varepsilon > 0$  we note from (3.4) that  $\tilde{r}(\varepsilon, 0) = \frac{1}{2}\lambda_1\varepsilon^2$  and so the solution to (3.3) with the choice  $p_0 = (\varepsilon, 0)$  is the curve  $\gamma_\beta$  with the angle  $\beta$  given by

$$\frac{1}{2} \left( \frac{\lambda_1 \varepsilon^2}{\lambda_1 + \lambda_2} \right) \cot \beta = A + O(\varepsilon),$$

cf. (3.5). Hence, by (3.7) we see that

$$d_{F_H}((0, 0, 0), (\varepsilon, 0, A)) = \sqrt{(A + O(\varepsilon))^2 (\lambda_1 + \lambda_2)^2 + \frac{1}{4} \lambda_1^2 \varepsilon^4}.$$

Sending  $\varepsilon \rightarrow 0$  and appealing to the continuity of the metric  $d_{F_H}$  we arrive at the result.  $\square$

## 4 Examples of non-existence

In light of Proposition 2.5, we see that a solution to (2.3) will exist if and only if a three-dimensional geodesic guaranteed by Theorem 2.4 contains no vertical segments over either of the wells. Often the vertical fiber over a well is not a three-dimensional geodesic and therefore none of its segments can be part of any geodesic. In this case Proposition 2.5 applies and the projection of the three-dimensional geodesic provides a solution to (2.3) for all values of the constraint. We give such an example at the end of the next sub-section. On the other hand, even if *both* vertical fibers over the wells *are* three-dimensional geodesics, solutions to (2.3) may still exist for all  $A$ -values. Such is the case of the homogeneous metrics considered in [1] and reviewed in Section 3.

In this section we present examples of explicit metrics and concrete areas  $A$  for which the problem (2.3) has no solution.

### Non-existence with one potential well

We begin with an example of non-existence for the case of  $F$  vanishing at only one point, which we take to be the origin in  $\mathbb{R}^2$ . We will then use this non-existence example to build an example of non-existence for the two well problem (2.3). For our one well example, we will take  $F^2$  to be a radial function  $F^2 = F^2(r)$ , with  $r = |p|$ . Specifically, we consider the case

$$F^2(r) = r^2 + br^4 \text{ for } b > 0. \quad (4.1)$$

Then for this  $F^2$  and for any  $p_0 \in \mathbb{R}^2 \setminus \{(0, 0)\}$  and  $A \in \mathbb{R}$  we consider the problem

$$\inf \int_a^b \sqrt{|\gamma|^2 + b|\gamma|^4} |\gamma'| \, dt, \quad (4.2)$$

where the infimum is taken over all locally Lipschitz continuous curves  $\gamma : [a, b] \rightarrow \mathbb{R}^2$  such that  $\gamma(a) = p_0$ ,  $\gamma(b) = (0, 0)$  and such that  $\tilde{\mathcal{A}}(\gamma) := \int_\gamma \omega_1 = \tilde{A}$  where

$$\omega_1 := \frac{1}{2} r^2 d\theta = -\frac{1}{2} (-p_2 dp_1 + p_1 dp_2)$$

and  $\tilde{A}$  is any given real number.

Here we have changed the constraint from the one used earlier in this paper, namely  $\mathcal{A}(\gamma) := \int_{\gamma} \omega_0 = A$  where  $\omega_0 := p_1 dp_2$ , to  $\tilde{\mathcal{A}}(\gamma) = \tilde{A}$  for convenience in working with polar coordinates, but we point out that since  $d\omega_1 = d\omega_0$ , the two constraints are equivalent in the following sense: If we let  $\ell_0$  denote the directed line segment from the origin to the point  $p_0$  then by Stokes Theorem for any curve  $\gamma$  joining  $p_0$  to the origin we have

$$\tilde{\mathcal{A}}(\gamma) - \mathcal{A}(\gamma) = \int_{\gamma} \omega_1 - \omega_0 = \int_{\ell_0} \omega_1 - \omega_0 = - \int_{\ell_0} \omega_0 =: C_0. \quad (4.3)$$

Consequently, specifying a constraint value  $\mathcal{A}(\gamma) = A$  is the same as specifying the constraint value  $\tilde{\mathcal{A}}(\gamma) = A + C_0$ .

Now we will argue that problem (4.2) has a solution if and only if

$$|\tilde{A}| \leq \frac{\sqrt{|p_0|}}{2\sqrt{b}}. \quad (4.4)$$

To this end, given any admissible curve  $\gamma = \gamma(t)$  mapping say  $[0, 1]$  into  $\mathbb{R}^2$  we introduce two new scalar dependent variables  $R(t)$  and  $\alpha(t)$  via

$$R(t) := |\gamma(t)|^2 \quad \text{and} \quad \alpha(t) := 4 \int_{\gamma([0,t])} \omega_1. \quad (4.5)$$

Then a routine calculation reveals the relation

$$\int_0^1 \sqrt{|\gamma|^2 + b|\gamma|^4} |\gamma'| dt = \int_0^1 \frac{\sqrt{1+bR}}{2} \sqrt{(R')^2 + (\alpha')^2} dt =: \tilde{E}(R, \alpha).$$

Thus, this change of variables “desingularizes” the problem in favor of one which is conformal to the standard Euclidean metric in the right-half  $R - \alpha$  plane with a strictly positive conformal factor. What is more, the constraint now simply becomes part of a Dirichlet condition,

$$(R(0), \alpha(0)) = (|p_0|, 0) \quad \text{and} \quad (R(1), \alpha(1)) = (0, 4\tilde{A}). \quad (4.6)$$

Of course there is also the crucial constraint  $R \geq 0$  and the equivalence of the constrained problem (4.2) and the minimization of  $\tilde{E}$  subject to (4.6) relies on the solution to the latter problem avoiding any line segment along the  $\alpha$ -axis where  $\omega_1$  is not well-defined, thereby making the variable  $\alpha$  no longer an accurate measure of area.. We should remark that this idea of introducing a variable to keep track of the constraint (area) value was already employed in our earlier reformulation of (2.3) in terms of three-dimensional geodesics, but the difference here is that due to the assumed radial dependence of the conformal factor, the resultant problem now is still two-dimensional.

The minimization of  $\tilde{E}(R, \alpha)$  subject to the Dirichlet conditions (4.6) and  $R \geq 0$  can be carried out explicitly. First, we observe that since the conformal factor is independent of  $\alpha$  and since it is a monotone increasing function of  $R$ , it is easy to argue that whenever  $R$  is positive, it is always optimal to consider competitors that are graphs over the positive  $R$  axis. In other words, it is always unnecessarily costly for a curve to double back when progressing from  $R = |p_0|$  to  $R = 0$ . What could happen, however, is that the minimizer includes a vertical line segment along

the  $\alpha$ -axis, that is along the line  $R = 0$ . This phenomenon is exactly the scenario where the three-dimensional minimizing curve of Theorem 2.4 does *not* equate to a solution to (2.3).

With this in mind, we now ask: What is the set of  $\tilde{A}$  values such that the minimizer to  $\tilde{E}(R, \alpha)$  is never vertical over the  $\alpha$ -axis and so remains a graph for all  $t \in [0, 1]$ ? Phrasing the minimization then for graphs  $\alpha = f(R)$ , we arrive at the problem

$$\inf_f \int_0^{|p_0|} \frac{\sqrt{1+bR}}{2} \sqrt{1+(f'(R))^2} dR \quad (4.7)$$

subject to the boundary conditions  $f(0) = 4\tilde{A}$  and  $f(|p_0|) = 0$ .

Criticality for this problem takes the form

$$\sqrt{1+bR} \frac{f'}{\sqrt{1+(f')^2}} = C_1$$

for a constant  $C_1$  and by considering the limit  $R \rightarrow 0$ , we see that evidently  $|C_1| \leq 1$ . Integrating this ODE we obtain a family of parabolas  $\{f_{C_1}\}$  opening to the right given by

$$f_{C_1}(R) = -\frac{2C_1}{b} \sqrt{1-C_1^2+bR} + D \quad (4.8)$$

with the minus sign arising if we consider, for example, the case  $\tilde{A} > 0$  and  $D$  is chosen so that  $f_{C_1}(|p_0|) = 0$ . The question we posed above (4.7) now takes the concrete form: What is the largest value of  $\tilde{A}$  such that one of these lower branches of parabolas makes the transition from  $4\tilde{A}$  down to zero on the  $R$  interval  $[0, |p_0|]$ ? After a little bit of algebra we arrive at the answer, namely (4.4), which corresponds to the parabola  $f_1$  with  $C_1 = 1$  whose vertex meets the  $\alpha$ -axis tangentially.

For any  $|\tilde{A}| > \frac{\sqrt{|p_0|}}{2\sqrt{b}}$ , the parabola satisfying (4.6) will bow into the left half-plane where  $R$  is negative. Such a curve is indeed length minimizing in the metric defined by  $\tilde{E}$  provided it stays in the half-plane  $R > -\frac{1}{b}$ , but it is inadmissible given the requirement  $R \geq 0$  forced by (4.5). Since the conformal factor is always cheaper along the  $\alpha$ -axis than in the region  $R > 0$ , the optimal resolution for such large  $\tilde{A}$  is clearly to join two points along the  $\alpha$ -axis with a vertical segment. (See Figure 1a.) Going back to the formulation (2.5), this provides an example of a three-dimensional minimizer from Theorem 2.4 whose projection fails to provide a solution to problem (4.2); hence (4.2) can have no solution.

We should remark that while for constraint values satisfying (4.4), the minimizer in the  $R\alpha$  plane is a piece of a parabola, when one reverts to standard polar coordinates, one finds the optimal curve in the  $p_1 p_2$  plane is given by an (infinite) spiral satisfying the relation

$$\frac{d\theta}{dr} = -\frac{C_1}{r\sqrt{1-C_1^2+br^2}} \quad \text{for } 0 < r \leq |p_0|. \quad (4.9)$$

Working with (4.9), one readily checks that these minimizing spirals satisfy the Euler-Lagrange equation (2.12) with Lagrange multiplier  $\lambda = 2C_1$  so that by our bound on  $C_1$  we see that

$$|\lambda| \leq 2. \quad (4.10)$$

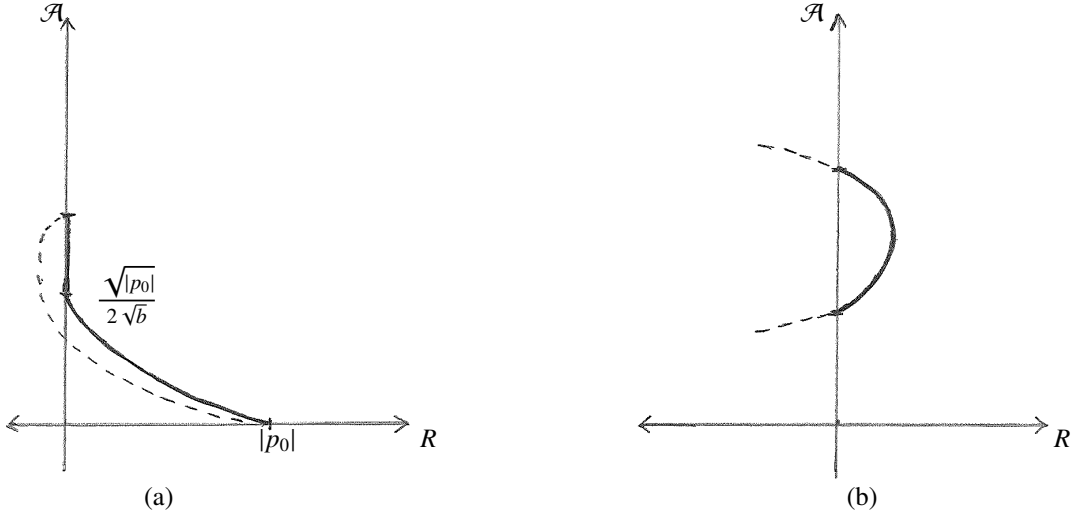


Figure 1: (a) The appearance of a vertical segment once the value of area exceeds a critical value for  $b > 0$ . The dashed parabola depicts the optimal curve if negative  $R$  values were allowable. (b) When  $b < 0$  no vertical segment is geodesic since the parabola is shorter.

This inequality will be used in the two-well example below.

A more interesting consequence of this formula is that when applied to the limiting case  $C_1 = 1$ , the derivative  $\frac{d\theta}{dr}$  behaves like  $Const./r^2$  thus showing that this  $E$ -minimizing spiral has *infinite* Euclidean arclength, in contrast for instance to the situation in the homogeneous metric, cf. (3.6).

Before moving on to an example of non-existence for the two-well problem, let us consider the situation where  $F^2$  is given by (4.1) in a neighborhood of the origin but with the coefficient  $b$  now taken to be *negative*. Then (4.8) reveals critical points of  $\tilde{E}$  consisting of parabolas opening to the left. We consider now two nearby points on the  $\alpha$ -axis. Being close together one knows there is a unique geodesic joining them in this metric. We conclude that the unique geodesic joining these two points is the left-opening parabola. In particular, it must be shorter than the vertical line segment joining them. See Figure 1b.

Phrased in terms of the metric  $d_F$  this means the vertical fiber over the origin is not a three-dimensional geodesic since that would force arbitrarily small vertical segments to be length-minimizing. Applying this to the two-well setting, we conclude from Proposition 2.5 that if  $F$  is

locally described by (4.1) near both of its wells with  $b < 0$ , then a solution exists to (2.3) for all values of the constraint  $A$ .

### Non-existence with two potential wells

We will now use the previous example to construct a potential  $F^2$  having two potential wells for which problem (2.3) has no solution. For a constant  $k > 1$ , later taken to be sufficiently large, we define the function

$$g(r) := \begin{cases} r^2 + (k^2 - 1)r^4 & \text{for } r \leq 1, \\ k^2 & \text{for } r \geq 1. \end{cases} \quad (4.11)$$

Then we take  $F^2 : \mathbb{R}^2 \rightarrow \mathbb{R}$  to be a given in the left half-plane by  $g$  in a polar coordinate system centered at  $(-1, 0)$  and in the right half-plane we take the even reflection of this function. The resulting function  $F^2$  vanishes at the points  $(\pm 1, 0)$ , is locally radial about these wells in discs which we denote by  $D_+$  and  $D_-$  of radius one centered at  $(\pm 1, 0)$  that touch at  $(0, 0)$ . Outside these discs,  $F$  is the constant  $k^2$ . While this  $F$  is only Lipschitz continuous across the boundary of the discs, this amount of smoothness will suffice for our purposes, though the example could be readily modified to provide  $C^2$  examples. We also observe that the unique unconstrained planar geodesic here is clearly the horizontal line segment joining the two wells so that in particular by (4.3) we have  $\tilde{\mathcal{A}}(\gamma) = \mathcal{A}(\gamma)$  for any curve joining the wells.

We will now argue that there is an interval of area values  $\mathcal{A}(\gamma) = A$ —actually two intervals if one allows for  $A$  both negative and positive—for which there is no solution to (2.3). First we note that the line segment joining the two wells is clearly the minimizing two-dimensional geodesic in the sense of (2.1) for this conformal factor  $F$ . Now we suppose  $\gamma$  is a solution to (2.3) for a fixed value of  $A$  which we may as well take to be positive, since the case of negative  $A$  is entirely the same.

We observe that within either  $D_+$  or  $D_-$ ,  $\gamma$  solves a one-well problem (4.2) as described in the previous example with the value of  $b$  given by  $k^2 - 1$ , some (unknown) value of  $A$  and the (unknown) point  $p_0$  satisfying  $|p_0| = 1$ . Since there is a global Lagrange multiplier associated with this problem, we see that  $\gamma$  must solve (2.12) in both discs and outside the two discs for the same value of  $\lambda$ . Hence within both discs, the solution must be a spiral given by (4.9) with the *same* constant  $C_1 = \lambda/2$  and as before  $|\lambda| \leq 2$ . Consequently, we conclude that within both discs the two spirals must agree up to a rotation. What is more, the amount of area  $\mathcal{A}(\gamma)$  must agree and so by (4.4) the total cannot exceed the value  $\frac{1}{\sqrt{k^2-1}}$ , which is small when  $k$  is large.

Now outside of the two discs, since the metric is just the constant  $k$  multiplying the standard Euclidean metric, we know that  $\gamma$  is a circular arc. Since  $\gamma$  is  $C^1$  across the boundary of the discs and since the spirals in the two discs agree up to a rotation, we conclude that the circular arc makes the same angle with the two circles forming the boundaries of these discs. Furthermore, from (4.9), this contact is nearly orthogonal when  $k$  is taken to be large. Then denoting the radius of curvature of this circular arc by say  $\bar{R}$ , one can easily check that the condition of constancy of the Lagrange multiplier within and without the discs implies through (4.10) that

$$\frac{2k}{\bar{R}} = \lambda \leq 2,$$

so that the curvature of this arc is large when  $k$  is taken to be large.

Elementary geometric considerations then lead to the conclusion that in fact  $\gamma$ , and in particular the circular arc, must be symmetric about the vertical  $p_2$  axis. For  $k$  large there are only two

possibilities: either the circular arc is close to the  $p_1$  axis, thereby contributing a small amount to the total value of  $\mathcal{A}(\gamma)$  or else it spans a huge amount of area in region  $p_2 > 0$ , see Figure 2.

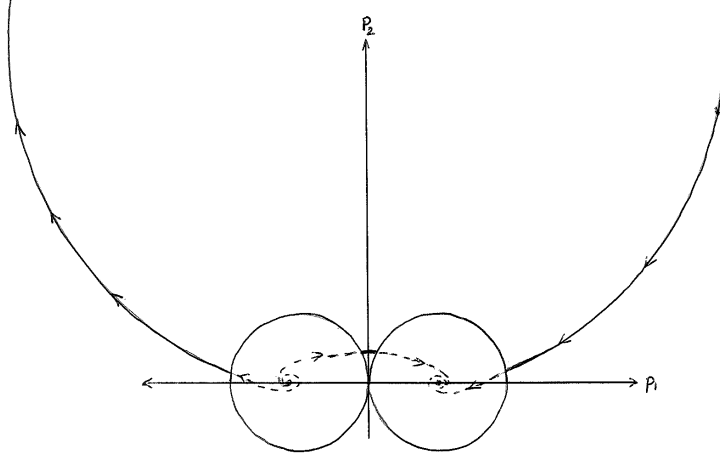


Figure 2: Counter-example for intermediate area values: When  $k$  is large, either a tiny circular arc bridges the two spirals or else a huge circular arc is required.

We conclude that the only attainable values of area  $\mathcal{A}$  for this example are either

$$\left(-\frac{1}{\sqrt{k^2-1}} - \delta, \frac{1}{\sqrt{k^2-1}} + \delta\right) \quad \text{for } \delta \ll 1 \quad \text{or} \quad |A| \geq M \quad \text{for some } M \gg 1,$$

so there is a large interval of positive  $A$ -values for which no solution exists.

Finally we remark one could likely produce a different type of example of non-existence by taking  $F$  in neighborhoods of the wells to again resemble our one-well non-existence example but then taking  $F$  to be huge outside these neighborhoods except for a narrow ‘trench’ along the line segment joining the wells. Such an example would presumably lead to non-existence for all constraint values exceeding some positive constant in absolute value.

## 5 Existence for constraint values near $\mathcal{A}(\gamma_0)$

In Section 2 we introduced a three-dimensional re-phrasing of our problem (2.3) and in Proposition 2.5 we noted that the three-dimensional geodesic has a projection solving (2.3) provided this geodesic does not contain any line segment over either of the wells  $\mathbf{p}_-$  or  $\mathbf{p}_+$ . In the previous section, however, we exhibited an example to show that such a vertical line segment can at times exist, leading to non-existence for (2.3). In this section we turn to our main result asserting that for constraint values near that of the unconstrained planar geodesic  $\gamma_0$  there is never a vertical line segment for the length-minimizing geodesic in three-dimensions. Hence, its projection in the plane will solve our problem.



Before getting into the details of the proof we want to describe the main idea behind the argument. For heuristic purposes, let us suppose that for an interval of area constraint values  $A$  we have a family of minimizers to our problem (2.3) smoothly depending on  $A$ , which we denote by  $\gamma_A$ . For each  $A$  then the curve  $\gamma_A$  will be a critical point satisfying the ODE (2.12) with corresponding Lagrange multiplier  $\lambda_A = F^2(\gamma_A)\kappa_g(\gamma_A)$ , cf. Remark 2.7. Computing the derivative of length  $E(\gamma_A)$  with respect to area  $A$  we find

$$\begin{aligned}\frac{dE(\gamma_A)}{dA} &= \frac{d}{dA} \int F(\gamma_A) |\gamma'_A| = \int \left[ - \left( F(\gamma_A) \frac{\gamma'_A}{|\gamma'_A|} \right)' + |\gamma'_A| \nabla F(\gamma_A) \right] \cdot \frac{\partial \gamma_A}{\partial A} \\ &= -\lambda_A \int (\gamma'_A)^\perp \cdot \frac{\partial \gamma_A}{\partial A}.\end{aligned}$$

Then differentiating the constraint condition  $\mathcal{A}(\gamma_A) = A$  with respect to  $A$  as well we see that

$$- \int (\gamma'_A)^\perp \cdot \frac{\partial \gamma_A}{\partial A} = 1 \quad \text{so that} \quad \frac{dE(\gamma_A)}{dA} = \lambda_A. \quad (5.1)$$

Now for the constraint value associated with the unconstrained planar geodesic  $\gamma_0$ , namely  $A = \mathcal{A}(\gamma_0)$ , we know  $\lambda = 0 = \kappa_g$  so we should expect that for nearby  $A$ -values,  $\lambda_A$  should be *small*. On the other hand, when a minimizing curve  $\gamma_A$  approaches a zero of  $F$  where  $F$  is well-approximated by its quadratic Taylor polynomial  $F_H$ , we might anticipate that the rate of change of length with respect to area—that is,  $\lambda_A$ —should resemble more and more that of the quadratic case for which we have a complete solution as described in Section 3. In particular, through Corollary 3.4 we know that the cost of a vertical line segment measured in length per area for the homogeneous metric  $d_{F_H}$  is  $\lambda = \lambda_1 + \lambda_2$ , which is ‘far’ from zero. Therefore when  $A$  is near  $\mathcal{A}(\gamma_0)$ , the three-dimensional geodesic cannot become vertical over either well.

We begin with a key lemma that says this rate of change for vertical segments measured with respect to our general metric  $d_F$  is given by the value inherited from  $d_{F_H}$ .

**Lemma 5.1.** *Translating either  $\mathbf{p}_+$  or  $\mathbf{p}_-$  to the origin, the value of the rate of change of length with respect to area for a vertical segment over either well is given by*

$$\lim_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (0, 0, \alpha))}{|\alpha|} = \lambda_1 + \lambda_2, \quad (5.2)$$

where  $F$  takes the form

$$F(p_1, p_2) = \sqrt{\lambda_1^2 p_1^2 + \lambda_2^2 p_2^2} + O(|p|^3) \quad \text{in a neighborhood of the origin} \quad (5.3)$$

for positive constants  $\lambda_1$  and  $\lambda_2$ .

**Remark 5.2.** *The pair of positive constants  $\lambda_1$  and  $\lambda_2$  for the Taylor development of  $F$  near  $\mathbf{p}_+$  is not required to be the same pair for  $F$  near  $\mathbf{p}_-$ .*

*Proof.* We will show this for  $\alpha \rightarrow 0^+$  with the result for  $\alpha \rightarrow 0^-$  following similarly. To see this, we invoke the triangle inequality

$$\frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} \leq \frac{d_F((0, 0, 0), (\alpha^2, 0, \alpha))}{\alpha} + \frac{d_F((\alpha^2, 0, \alpha), (0, 0, \alpha))}{\alpha}$$

and for the last term we have

$$\frac{d_F((\alpha^2, 0, \alpha), (0, 0, \alpha))}{\alpha} \leq \frac{\int_0^{\alpha^2} F(t, 0) dt}{\alpha} \rightarrow 0.$$

Now let  $\gamma_{\beta(\alpha)}$  denote the projection of the optimal three-dimensional curve in the homogeneous metric  $d_{F_H}$ , that is, the planar curve joining  $(0, 0)$  to  $(\alpha^2, 0)$  and satisfying the constraint value  $\mathcal{A} = \alpha$  as described in Theorem 3.1. In particular,  $\beta(\alpha)$  is given by (3.5) with  $p_0 = (\alpha^2, 0)$  and  $A = \alpha$ . Since  $|\gamma_{\beta(\alpha)}| \leq \alpha^2$ , we are working near the origin so we can use the estimate

$$|F(p) - F_H(p)| \leq C |p|^2,$$

cf. (3.2) and apply (3.6) to argue that

$$\begin{aligned} \limsup_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} &\leq \limsup_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (\alpha^2, 0, \alpha))}{\alpha} \leq \limsup_{\alpha \rightarrow 0} \frac{\int F(\gamma_{\beta(\alpha)}) |\gamma'_{\beta(\alpha)}|}{\alpha} \\ &\leq \limsup_{\alpha \rightarrow 0} \left\{ \frac{\int F_H(\gamma_{\beta(\alpha)}) |\gamma'_{\beta(\alpha)}|}{\alpha} + \frac{1}{\alpha} \int |F(\gamma_{\beta(\alpha)}) - F_H(\gamma_{\beta(\alpha)})| |\gamma'_{\beta(\alpha)}| \right\} \\ &\leq \limsup_{\alpha \rightarrow 0} \left\{ \frac{d_{F_H}((0, 0, 0), (\alpha^2, 0, \alpha))}{\alpha} + \frac{1}{\alpha} C \alpha^2 \int |\gamma'_{\beta(\alpha)}| \right\} \\ &= \lim_{\alpha \rightarrow 0} \frac{d_{F_H}((0, 0, 0), (\alpha^2, 0, \alpha))}{\alpha}. \end{aligned}$$

As (3.7) gives us an explicit formula for  $d_{F_H}((0, 0, 0), (\alpha^2, 0, \alpha))$ , we conclude that

$$\limsup_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} \leq \lim_{\alpha \rightarrow 0} \frac{\sqrt{\alpha^2(\lambda_1 + \lambda_2)^2 + \frac{1}{4}\lambda_1^2\alpha^4}}{\alpha} = \lambda_1 + \lambda_2. \quad (5.4)$$

To complete the demonstration of (5.2) we must show the reverse inequality. For this purpose we introduce the scaled conformal factor  $F_\theta(p) := \frac{1}{\theta} F(\theta p)$  for any  $\theta > 0$  and the corresponding functional

$$E_\theta(\zeta) := \int F_\theta(\zeta) |\zeta'|$$

for any planar curve  $\zeta$ . We note that  $F_\theta \rightarrow F_H$  as  $\theta \rightarrow 0$  uniformly on compact subsets of  $\mathbb{R}^2$  so that  $d_{F_\theta}(P, Q) \rightarrow d_{F_H}(P, Q)$  for all  $P$  and  $Q$  in  $\mathbb{R}^3$  as well. Here  $d_{F_\theta}$  denotes the metric as defined in (2.5) but with the conformal factor taken to be  $F_\theta$ . Now for any  $\alpha$  let  $\gamma_\alpha : [a, b] \rightarrow \mathbb{R}^2$  be any curve with endpoints  $\gamma_\alpha(a) = \gamma_\alpha(b) = (0, 0)$  such that  $\mathcal{A}(\gamma_\alpha) = \alpha$  and such that

$$E(\gamma_\alpha) \leq d_F((0, 0, 0), (0, 0, \alpha)) + \alpha^2;$$

that is,  $\gamma_\alpha$  is a good competitor in the minimization of  $E$  subject to these endpoint and constraint values. We then observe that

$$\begin{aligned} d_{F_{\sqrt{\alpha}}}((0, 0, 0), (0, 0, 1)) &\leq E_{\sqrt{\alpha}}\left(\frac{1}{\sqrt{\alpha}}\gamma_\alpha\right) \quad \left(\text{since } \mathcal{A}\left(\frac{1}{\sqrt{\alpha}}\gamma_\alpha\right) = 1\right) \\ &= \frac{1}{\alpha} E(\gamma_\alpha) \leq \frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} + \alpha. \end{aligned}$$

Letting  $\alpha \rightarrow 0$  and invoking Corollary 3.4 we arrive at the desired reverse inequality

$$\liminf_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} \geq d_{F_H}((0, 0, 0), (0, 0, 1)) = \lambda_1 + \lambda_2$$

and so (5.2) is established.  $\square$

Now we can establish our main result.

**Theorem 5.3.** *Let  $\gamma_0 : [0, 1] \rightarrow \mathbb{R}^2$  be an unconstrained minimizing geodesic joining  $\mathbf{p}_-$  to  $\mathbf{p}_+$ , i.e. a solution to the variational problem (2.1) with  $p = \mathbf{p}_-$  and  $q = \mathbf{p}_+$ . Then there exists a number  $\varepsilon_0 > 0$  such that for all  $A$  in the interval  $(\mathcal{A}(\gamma_0) - \varepsilon_0, \mathcal{A}(\gamma_0) + \varepsilon_0)$ , there exists a solution to (2.3).*

*Proof.* For any  $\varepsilon \in \mathbb{R}$ , later to be taken small in absolute value, let  $\Gamma_\varepsilon : [0, 1] \rightarrow \mathbb{R}^3$  be the length-minimizing geodesic guaranteed by Theorem 2.4 satisfying  $L(\Gamma_\varepsilon) = d(P_\varepsilon, Q_\varepsilon)$  where  $P_\varepsilon$  and  $Q_\varepsilon$  are any two points in  $\mathbb{R}^3$  such that

$$\Pi(P_\varepsilon) = \mathbf{p}_-, \quad \Pi(Q_\varepsilon) = \mathbf{p}_+ \quad \text{and} \quad Q_\varepsilon^{(3)} - P_\varepsilon^{(3)} = A_0 + \varepsilon. \quad (5.5)$$

Throughout the proof we will denote by  $\gamma_\varepsilon$  the projection  $\Pi(\Gamma_\varepsilon)$ .

In light of Theorem 2.5 it will suffice to demonstrate that  $\Gamma_\varepsilon$  does not possess any vertical line segment over either  $\mathbf{p}_-$  or  $\mathbf{p}_+$ . To this end, we will pursue an argument by contradiction and suppose that this condition *fails* for a sequence  $\varepsilon_j \rightarrow 0$ . For it to fail near  $\mathbf{p}_-$  means that we assume:

For a sequence  $\varepsilon_j \rightarrow 0$  there is a nontrivial interval  $I_j \subset [0, 1]$  such that  $\gamma_{\varepsilon_j}(t) = \mathbf{p}_-$  for all  $t \in I_j$ . (5.6)

The argument precluding such a scenario from happening near  $\mathbf{p}_+$  is identical.

After an appropriate translation of coordinates in the  $p_1 p_2$ -plane we may take  $\mathbf{p}_- = (0, 0)$  and after an appropriate rotation of coordinates so that the coordinate axes are aligned with the eigenvectors of the Hessian  $D^2 W(\mathbf{p}_-)$  we may assume the function  $F = \sqrt{W}$  near  $\mathbf{p}_-$  takes the form given by (5.3).

Our strategy will be to first construct a family of competitors smoothly depending on a parameter, say  $\alpha$ , that involves either enlarging or shrinking the length of the assumed line segment within  $\Gamma_{\varepsilon_j}$  by an amount  $\alpha$  while compensating for this change by subtracting or adding a “bump” to  $\Gamma_{\varepsilon_j}$  away from the segment so as to maintain the endpoint condition (5.5). Since for this portion of the argument, the value of  $\varepsilon_j$  is not varying, for ease of notation we will temporarily suppress this dependence and write simply  $\Gamma$  for  $\Gamma_{\varepsilon_j}$ . Once we need to consider  $\Gamma_{\varepsilon_j}$  for  $\varepsilon_j$  small we will revert to the original notation.

For any small  $\alpha \in \mathbb{R}$  let us describe more precisely this smooth deformation of  $\Gamma$ , which we denote by  $\Gamma^\alpha$ , as follows. We alter the vertical line segment  $\Gamma(I_j)$  by length  $\alpha$ . Recall that the third component of  $\Gamma$  effectively keeps track of the value of the area constraint  $\mathcal{A}$ . Therefore, in order to create a family of deformations that preserves the total change in the third component, we work locally near some planar sub-arc of  $\gamma := \Pi(\Gamma)$  away from  $\mathbf{p}_-$  and  $\mathbf{p}_+$ , say  $\gamma|_{[a,b]}$  for some parameter interval  $[a, b]$  and simultaneously modify this sub-arc by augmenting it with a bump that subtracts  $\alpha$  from the constraint value  $\mathcal{A}(\gamma)$ . For example, we may suppose that on the interval

$[a, b]$ , this sub-arc is expressible as a graph  $y = f(x)$  of a smooth scalar function  $f$ . Then letting  $g : [a, b] \rightarrow \mathbb{R}$  be any smooth function satisfying the conditions

$$g(a) = 0 = g(b) \quad \text{and} \quad \int_a^b g(x) dx = 1,$$

we consider the family of perturbations  $\tilde{\gamma}_\alpha : [a, b] \rightarrow \mathbb{R}^2$  of  $\gamma_{|[a,b]}$  given by  $x \mapsto (x, f(x) + \alpha g(x))$ . Note that

$$\mathcal{A}(\tilde{\gamma}_\alpha) := \int_a^b \tilde{\gamma}_\alpha^{(1)} \tilde{\gamma}_\alpha^{(2)'} dx = \int_a^b x(f'(x) + \alpha g'(x)) dx = \mathcal{A}(\gamma_{|[a,b]}) + \alpha \int_a^b x g'(x) dx = \mathcal{A}(\gamma_{|[a,b]}) - \alpha.$$

Along this parameter interval we then define the third component of  $\Gamma^\alpha$  through requirement (2.4).

As  $\Gamma^0 = \Gamma$  the family  $\{\Gamma^\alpha\}$  constitutes a smooth deformation of the minimizing geodesic  $\Gamma$  that preserves its endpoints  $P_\varepsilon$  and  $Q_\varepsilon$  up to translation in the third coordinate direction, cf. (5.5). Invoking Lemma 5.1 along with (2.12) of Proposition 2.5 we can compute

$$\begin{aligned} \frac{dL(\Gamma^\alpha)}{d\alpha} \Big|_{\alpha=0} &= \lim_{\alpha \rightarrow 0} \frac{d_F((0, 0, 0), (0, 0, \alpha))}{\alpha} + \frac{d}{d\alpha} \Big|_{\alpha=0} \left\{ \int_a^b F(x, f(x) + \alpha g(x)) \sqrt{1 + (f' + \alpha g')^2} dx \right\} \\ &= \lambda_1 + \lambda_2 + \int_a^b \left[ - \left( F(\gamma) \frac{\gamma'}{|\gamma'|} \right)' + |\gamma'| |\nabla F(\gamma)| \cdot (0, g) \right] dx \\ &= \lambda_1 + \lambda_2 - \lambda \int_a^b (\gamma')^\perp \cdot (0, g) dx = \lambda_1 + \lambda_2 - \lambda \int_a^b g(x) dx = \lambda_1 + \lambda_2 - \lambda, \end{aligned}$$

where  $\lambda$  is the Lagrange multiplier arising in the Euler-Lagrange equation. Since  $\Gamma$  is minimizing, necessarily the derivative above must vanish and so we conclude that a minimizer that includes a vertical segment must have Lagrange multiplier  $\lambda$  given by

$$\lambda = \lambda_1 + \lambda_2. \quad (5.7)$$

We recall now that really  $\Gamma = \Gamma_{\varepsilon_j}$ , the length-minimizing geodesic subject to the  $\varepsilon$ -dependent conditions (5.5). Yet under the contradiction hypothesis (5.6) that  $\Gamma_{\varepsilon_j}$  includes a vertical line segment, the value of  $\lambda$  is forced to satisfy condition (5.7) which is independent of  $\varepsilon$ . We will now exploit this property to show that  $\inf L(\Gamma_{\varepsilon_j}) > E(\gamma_0)$  and easily reach a contradiction.

To this end, we recall the definition of the planar metric  $\bar{d}_F(p, q)$  given in (2.1) and we fix any positive  $r < \frac{1}{2} \bar{d}_F(\mathbf{p}_-, \mathbf{p}_+) = \frac{1}{2} E(\gamma_0)$ . Then we introduce the topological annulus in the plane

$$\mathcal{D}_r := \{p \in \mathbb{R}^2 : r < \bar{d}_F(\mathbf{p}_-, p) < 2r\}. \quad (5.8)$$

We note that since each  $\gamma_{\varepsilon_j}$  starts at  $\mathbf{p}_-$  and ends at  $\mathbf{p}_+$ , it must contain at least one sub-arc lying inside  $\mathcal{D}_r$  with endpoints  $p_j$  and  $q_j$  satisfying  $\bar{d}_F(\mathbf{p}_-, p_j) = r$  and  $\bar{d}_F(\mathbf{p}_-, q_j) = 2r$  respectively. Denoting this sub-arc by  $\zeta_j$  and parametrizing it so that

$$F(\zeta_j(\ell)) |\zeta_j'(\ell)| = E(\zeta_j) \quad (5.9)$$

we obtain a family of curves  $\{\zeta_j\}$  mapping  $[0, 1]$  to  $\overline{\mathcal{D}_r}$ . Also we observe that  $r \leq E(\zeta_j)$  since the width in the metric  $\bar{d}_F$  of the annulus is  $r$ , while

$$\sup_j E(\zeta_j) < \sup_j E(\gamma_{\varepsilon_j}) \leq \sup_{\varepsilon \in [-1, 1]} d_F((\mathbf{p}_-, 0), (\mathbf{p}_+, \mathcal{A}(\gamma_0) + \varepsilon)) < \infty.$$

We will now use these  $j$ -independent bounds as well as (5.7) to establish the claim that

$$\liminf_{j \rightarrow \infty} E(\zeta_j) > r. \quad (5.10)$$

We recall that under contradiction hypothesis (5.6) we have shown that necessarily  $\gamma_{\varepsilon_j}$  solves (2.12) with  $\lambda$  given by (5.7). Then we use the chosen parametrization (5.9) to see that  $\zeta_j$  solves the system of ODE's

$$-\frac{1}{E(\zeta_j)} \left( F^2(\zeta_j) \zeta_j' \right)' + \frac{E(\zeta_j)}{F(\zeta_j)} \nabla F(\zeta_j) = -(\lambda_1 + \lambda_2) (\zeta_j')^\perp. \quad (5.11)$$

Since  $F$  is  $C^1$  and

$$0 < \min_{p \in \bar{D}_r} F(p) < \max_{p \in \bar{D}_r} F(p) < \infty,$$

being a solution to this system implies a  $C^2$  bound on the curves  $\{\zeta_j\}$  that is independent of  $j$ . Hence, we obtain a curve  $\zeta_0$  that is the uniform  $C^1$  limit of a subsequence of  $\{\zeta_j\}$  and  $\zeta_0$  must solve

$$-\frac{1}{E(\zeta_0)} \left( F^2(\zeta_0) \zeta_0' \right)' + \frac{E(\zeta_0)}{F(\zeta_0)} \nabla F(\zeta_0) = -(\lambda_1 + \lambda_2) (\zeta_0')^\perp \quad (5.12)$$

(first weakly, but then by standard regularity theory, strongly). Then the claim (5.10) follows because were it false, we would have  $E(\zeta_0) = r$ , making  $\zeta_0$  an  $E$ -minimizing curve spanning  $\mathcal{D}_r$  which would force  $\zeta_0$  to solve the above ODE with zero right-hand side.

With (5.10) in hand, we now have that for all  $\varepsilon_j$  sufficiently small:

$$L(\Gamma_{\varepsilon_j}) - E(\gamma_0) \geq E(\gamma_{\varepsilon_j}) - E(\gamma_0) > \frac{1}{2} \left( \liminf_{j \rightarrow \infty} E(\zeta_j) - r \right) =: a_0.$$

Clearly by modifying  $\gamma_0$ , say with a small bump, we can accommodate the constraint requirement  $\mathcal{A} = \mathcal{A}(\gamma_0) + \varepsilon_j$  for  $\varepsilon_j$  small to build a competing planar curve—and then three-dimensional curve via (2.4)—whose extra length is less than  $a_0$ , thus contradicting the minimality of  $\Gamma_{\varepsilon_j}$  among curves joining  $(\mathbf{p}_-, 0)$  to  $(\mathbf{p}_+, \mathcal{A}(\gamma_0) + \varepsilon_j)$  under assumption (5.6). Since no vertical line segment exists, we conclude from Proposition 2.5 that  $\gamma_{\varepsilon}$  solves (2.3) provided  $|\varepsilon|$  is sufficiently small.  $\square$

Next we establish that for values of  $A$  near  $\mathcal{A}(\gamma_0)$  the solutions from Theorem 5.3 contain no bubbles. This property will in particular be important when interpreting these curves as traveling waves to the Hamiltonian system discussed in Section 6.

**Theorem 5.4.** *Denote by  $\gamma_{\varepsilon}$  the solution to (2.3) obtained in Theorem 5.3 for  $A = \mathcal{A}(\gamma_0) + \varepsilon$  and  $|\varepsilon| < \varepsilon_0$ . Then there exists a positive value  $\varepsilon_1 \leq \varepsilon_0$  such that when  $|\varepsilon| < \varepsilon_1$  the curve  $\gamma_{\varepsilon}$  only meets  $\mathbf{p}_+$  and  $\mathbf{p}_-$  at its endpoints.*

*Proof.* We again proceed by way of contradiction. Thus we assume that for some sequence  $\varepsilon_j \rightarrow 0$ :

$$\begin{aligned} &\text{There exist two sequences } 0 \leq s_{\varepsilon_j} < t_{\varepsilon_j} < 1 \text{ such that} \\ &\gamma_{\varepsilon_j}(s_{\varepsilon_j}) = \gamma_{\varepsilon_j}(t_{\varepsilon_j}) = \mathbf{p}_- \text{ while } \gamma_{\varepsilon_j}(t) \notin \{\mathbf{p}_-, \mathbf{p}_+\} \text{ for } s_{\varepsilon_j} < t < t_{\varepsilon_j}. \end{aligned} \quad (5.13)$$

We will reach a contradiction under the scenario (5.13) in which, say,  $\ell_j := \gamma_{\varepsilon_j}([s_j, t_j])$  constitutes a closed planar loop (i.e. a ‘bubble’) passing through  $\mathbf{p}_-$ , by a line of reasoning similar to that employed in the existence proof of Theorem 5.3.

If we let  $q_j \in \ell_j$  denote the furthest point on the loop from  $\mathbf{p}_-$  in the  $d_F$  metric, i.e.

$$d_F(q_j, \mathbf{p}_-) = \max_{p \in \ell_j} d_F(p, \mathbf{p}_-),$$

we first argue that necessarily

$$d_F(q_j, \mathbf{p}_-) \rightarrow 0 \text{ as } j \rightarrow 0. \quad (5.14)$$

Otherwise there would exist a subsequence (still denoted by  $\varepsilon_j$ ) and a positive number  $a_0$  such that along this subsequence the length  $E(\gamma_{\varepsilon_j})$  would exceed the value  $E(\gamma_0)$  by at least  $a_0$  and the addition of a small bump to  $\gamma_0$  would create a competitor that beats  $\gamma_{\varepsilon_j}$ .

We then consider the smallest positive number  $C_j$  such that

$$\ell_j \subset \{p \in \mathbb{R}^2 : \frac{1}{2}\lambda_1 p_1^2 + \lambda_2 p_2^2 \leq C_j\},$$

with (5.14) now implying that  $C_j \rightarrow 0$ . Recalling Corollary 3.3, we observe that the elliptical boundary of this set is in fact a closed isoperimetric curve in the metric  $d_{F_H}$  corresponding to angle  $\beta = 0$ . Consequently it is critical for the constrained minimization of  $E_H$  and so the ellipse must satisfy (2.12) with conformal factor  $F_H(p) := \sqrt{\lambda_1^2 p_1^2 + \lambda_2^2 p_2^2}$  for some constant  $\lambda$ . In fact, in the case of minimizing ellipses enclosing area  $A$ , we are in a setting where the (explicit) solution clearly varies smoothly with  $A$  so that the heuristics presented at the outset of this section are in fact rigorous. Invoking (5.1) and the conclusion of Corollary 3.3 we then can identify the value of  $\lambda$  as  $\lambda_1 + \lambda_2$ . Note in particular that the Lagrange multiplier is independent of  $j$ . Expressing the ellipse as a parametric curve  $g_j : [a, b] \rightarrow \mathbb{R}^2$  it necessarily satisfies (2.12) and so when we take the inner product of this system of ODE’s with the vector  $(g'_j)^\perp$  we obtain the relation

$$F(g_j)\kappa(g_j) - \frac{1}{|g'_j|} \nabla F(g_j) \cdot (g'_j)^\perp = \lambda_1 + \lambda_2, \quad (5.15)$$

where we have introduced notation for the curvature of  $g_j$  with respect to the Euclidean metric, i.e.

$$\kappa(g_j) := \frac{1}{|g'_j|^3} g''_j \cdot (g'_j)^\perp.$$

Now each of the closed loops  $\ell_j$  is of course also a critical point solving (2.12) with its own corresponding Lagrange multiplier, say  $\lambda_j$ , so by a similar manipulation one has

$$F(\gamma_{\varepsilon_j})\kappa(\gamma_{\varepsilon_j}) - \frac{1}{|\gamma'_{\varepsilon_j}|} \nabla F(\gamma_{\varepsilon_j}) \cdot (\gamma'_{\varepsilon_j})^\perp = \lambda_j, \quad (5.16)$$

Since the contact between the ellipse  $g_j$  and the loop  $\gamma_{\varepsilon_j}$  must be tangential at  $q_j$  with the loop lying inside the ellipse, it follows that at  $q_j$  we have the relations

$$g_j = \gamma_{\varepsilon_j}, \quad \frac{(g'_j)^\perp}{|g'_j|} = \frac{(\gamma'_{\varepsilon_j})^\perp}{|\gamma'_{\varepsilon_j}|} \quad \text{and} \quad \kappa(\gamma_{\varepsilon_j}) \geq \kappa(g_j).$$

Therefore, subtracting (5.15) from (5.16) we conclude that

$$\lambda_j \geq \lambda_1 + \lambda_2. \quad (5.17)$$

Now if in addition we have an upper bound on  $\{\lambda_j\}$  that is independent of  $j$  then the contradiction is reached via compactness of  $\{\gamma_{\varepsilon_j}\}$  exactly as was done for the case of (5.6) where now (5.17) is used in place of (5.7).

Lastly, there remains the possibility that  $\lambda_j \rightarrow \infty$ . In this case, we again focus on an arc, say  $\zeta_j$ , of  $\gamma_{\varepsilon_j}$  that must span the topological annulus  $\mathcal{D}_r$  defined in (5.8). Working in a parametrization by Euclidean arclength so that  $|\zeta'_j| = 1$  we then conclude via (5.16) that necessarily the (Euclidean) curvature  $\kappa(\zeta_j)$  at every point of  $\zeta_j$  must tend to infinity with  $\lambda_j$ . Consequently, the Euclidean arclength of  $\gamma_{\varepsilon_j}$  in the annulus becomes infinitely long as well. Since in  $\mathcal{D}_r$ , the conformal factor  $F(\zeta_j)$  is bounded uniformly away from zero, this forces the value of  $E(\gamma_{\varepsilon_j})$  to infinity, obviously contradicting its minimality.  $\square$

Theorem 5.4 rules out the appearance of bubbles for the solutions obtained via Theorem 5.3 when the constraint value is sufficiently close to  $\mathcal{A}(\gamma_0)$ . For particular choices of the conformal factor  $F$ , however, such as the homogeneous choice of Section 3, we know that solutions may exist for large values of  $A$  as well. It seems quite likely that one can find examples of  $F$  for which solutions possessing bubbles exist at large values of the constraint.

## 6 Application to bi-stable Hamiltonian traveling waves

We conclude by recalling one of our motivations in considering the degenerate isoperimetric problem (2.3): the construction of traveling wave solutions to the bi-stable Hamiltonian system

$$\mathbb{J}u_t = \Delta u - \nabla W_u(u) \quad \text{for } u : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^2 \quad (6.1)$$

where  $\mathbb{J}$  denotes the symplectic matrix given by

$$\mathbb{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Here  $W : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a potential such as the one considered in this article but whereas the results from the previous sections apply with an arbitrary numbers of zeros, here we will insist that  $W$  vanishes at exactly two points. By a traveling wave solution to (6.1) we mean a solution taking the form

$$u(x, t) = U(x_1 - vt) = U(y) \quad \text{for some wave speed } v \in \mathbb{R}$$

that joins the two minima of  $W$ , that is, where  $U(\pm\infty) = \mathbf{p}_{\pm}$ . Such a  $U : \mathbb{R} \rightarrow \mathbb{R}^2$  must satisfy the system of ODE's

$$-v\mathbb{J}U' = U'' - \nabla_u W(U) \quad \text{for } -\infty < y < \infty, \quad U(\pm\infty) = \mathbf{p}_{\pm}. \quad (6.2)$$

One checks that (6.2) is the criticality condition associated with the constrained variational problem

$$\inf_{\mathcal{G}_A} H(u) \quad \text{where} \quad H(u) := \int_{-\infty}^{\infty} \frac{1}{2} |u'|^2 + W(u) dy \quad (6.3)$$

and where for a given  $A \in \mathbb{R}$  the admissible set  $\mathcal{G}_A$  is defined by

$$\mathcal{G}_A := \left\{ u : \mathbb{R} \rightarrow \mathbb{R}^2 : u - g \in H^1(\mathbb{R}; \mathbb{R}^2), \mathcal{A}(u) = A \right\}.$$

In the above definition,  $g : \mathbb{R} \rightarrow \mathbb{R}^2$  is any smooth function such that  $g(y) \equiv \mathbf{p}_+$  for  $y$  sufficiently large and  $g(y) \equiv \mathbf{p}_-$  for  $y$  sufficiently negative. The wave speed  $v$  arises as the Lagrange multiplier associated with the constraint  $\mathcal{A}(u) = A$  in making this connection.

Applying the trivial inequality  $H(u) \geq \sqrt{2}E(u)$  for  $E$  given by (2.2) and exploiting the convenient “equi-partition” parametrization  $|\gamma'| = \sqrt{2W(\gamma)}$  in competitors for (2.3) one can establish that any isoperimetric curve  $\gamma$  solving (2.3) is automatically a minimizer of (6.3), hence a solution to (6.2), provided this minimizer has no ‘bubbles’; that is, provided that scenario (5.13) does not occur for either  $\mathbf{p}_-$  or  $\mathbf{p}_+$ . This is the content of [1], Theorem 4.2. In bridging these two seemingly disparate problems the wave speed  $v$  is precisely the Lagrange multiplier  $\lambda = \lambda(A)$  associated with the minimizer  $\gamma$ . In light of Remark 2.7 this also means the wave speed can be characterized by the relation  $v = F^2(\gamma)\kappa_g$  where  $\kappa_g$  is the geodesic curvature of  $\gamma$ .

In [1] solutions to (2.3) without bubbles are obtained for constraint values near  $\mathcal{A}(\gamma_0)$  under the assumption that the potential  $W$  is identically equal to a non-degenerate quadratic in neighborhoods of  $\mathbf{p}_-$  and  $\mathbf{p}_+$ . In light of our existence result in Theorem 5.3 as well as Theorem 5.4 that rules out bubbles, we can now assert the corresponding result for more general potentials  $W$ :

**Theorem 6.1.** *Assume as in Theorem 5.3 that  $W : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a smooth function vanishing only at points  $\mathbf{p}_-$  and  $\mathbf{p}_+$  in  $\mathbb{R}^2$  with  $D^2W(\mathbf{p}_\pm) > 0$ . Let  $\gamma_0$  denote an unconstrained planar geodesic joining  $\mathbf{p}_-$  to  $\mathbf{p}_+$  as in Theorem 2.1. Then there exists a value  $\varepsilon_1 > 0$  such that for all  $A$  in the interval  $(\mathcal{A}(\gamma_0) - \varepsilon_1, \mathcal{A}(\gamma_0) + \varepsilon_1)$ , the curves, say  $\gamma_A$ , described in Theorem 5.3 and given an equipartition parametrization, will solve (6.3) and therefore solve (6.2) as well with a wave speed  $v_A$  depending on  $A$ .*

As noted in [1], without further assumptions there remains the possibility of a degeneracy of the curve  $\gamma_0$ , parametrized so that  $|\gamma'_0| = \sqrt{2W(\gamma_0)}$ , with respect to the energy  $H$  which could lead to the non-generic predicament that  $v_A = 0$  for all  $A$  in the interval of existence. In this situation the traveling waves would all in fact be standing waves, that is, simply a one-parameter family of heteroclinic connections solving

$$U'' - \nabla_u W(U) = 0 \quad \text{for } -\infty < y < \infty, \quad U(\pm\infty) = \mathbf{p}_\pm. \quad (6.4)$$

One way of eliminating this scenario is to add a non-degeneracy assumption regarding the second variation operator

$$\delta^2 H(\gamma)[\Phi] := \int_{-\infty}^{\infty} \left[ |\Phi'|^2 + \Phi \cdot D^2 W(U_0) \Phi \right] dy,$$

acting on  $\Phi \in H^1(\mathbb{R}; \mathbb{R}^2)$ . Here we quote Proposition 4.4 from [1]:

**Proposition 6.2.** *In addition to the assumptions of Theorem 6.1, assume that there exist constants  $R_0, c_0 > 0$  such that*

$$\nabla W(p) \cdot p \geq c_0 |p|^2 \quad \text{for all } p \in \mathbb{R}^2 \text{ with } |p| \geq R_0.$$

*Assume also that for every solution  $u_0$  to (6.4) that minimizes  $H$  among competitors satisfying  $u(\pm\infty) = \mathbf{p}_\pm$  one has that zero is an isolated simple eigenvalue of  $\delta^2 H(u_0)$ . Then, for any such  $H$ -minimizing heteroclinic, there exists  $\varepsilon_2 > 0$  such that if  $A \in (\mathcal{A}(\gamma_0) - \varepsilon_2, \mathcal{A}(\gamma_0) + \varepsilon_2)$  with  $A \neq \mathcal{A}(u_0)$  and if  $u_A : \mathbb{R} \rightarrow \mathbb{R}^2$  solves (6.3), then  $u_A$  solves (6.2) with wave speed  $v = v_A \neq 0$ .*



We point out that the function  $u'_0$  will always be an eigenfunction of  $\delta^2 H(u_0)$  with corresponding eigenvalue zero but this proposition asserts that if  $u'_0$  is the only such eigenfunction and if zero is isolated in the spectrum for any  $H$ -minimizing heteroclinic, then the curves produced in Theorem 5.3, appropriately parametrized, will constitute traveling waves that definitely travel.

## References

- [1] S. Alama, L. Bronsard, A. Contreras, J. Dadok and P. Sternberg, *A degenerate isoperimetric problem and traveling waves to a bi-stable Hamiltonian system*, to appear in C.P.A.M.
- [2] M.R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer Science & Business Media, 1999.
- [3] C. Carroll, A. Jacob, C. Quinn and R. Walters, *The isoperimetric problem on planes with density*, Bull. Austr. Math. Soc., **78**, (2008), 177-197.
- [4] A. Cañete, M. Miranda Jr. and D. Vittone, *Some isoperimetric problems in planes with density*, J. Geom. Anal., **20**, no. 2, (2010), 243-290.
- [5] J. Dahlberg, A. Dubbs, E. Newkirk and H. Tran, *Isoperimetric regions in the plane with density  $r^p$* , New York J. Math., **15**, (2010), 31-51.
- [6] M. Gromov, *Carnot-Carathéodory spaces seen from within*, Progress in Mathematics, **61**, Birkhäuser Boston, Inc., (1985).
- [7] S. Howe, *The log-convex density conjecture and vertical surface area in warped products*, Adv. in Geom., **15**, no. 4, (2015), 455-468.
- [8] A. Monteil and F. Santambrogio, *Metric methods for heteroclinic connections*, arXiv:1602.05487.
- [9] C. Rosales, A. Cañete, V. Bayle and F. Morgan, *On the isoperimetric problem in Euclidean space with density*, Calc. Var. PDE, **31**, (2008), 27-46.
- [10] A. Zuñiga and P. Sternberg, *On the heteroclinic connection problem for multi-well gradient systems*, to appear in J.D.E.